6204641333

**1.** From the data for 46 states in the United States for 1992, results of the regression are displayed as follows.

$$\ln C_i = 4.30 - 1.34 \ln P_i + 0.17 \ln Y_i$$
$$se \quad = (0.91) \quad (0.32) \qquad (0.20)$$
$$\bar{R}^2 = 0.27$$

where $C_i$ = cigarette consumption, packs per year

$P_i$ = real price per pack, \$ per pack

$Y_i$ = real disposable income per capita, \$ per week

1.1) Do the estimation results follow the law of demand?

Yes, because real disposable income per capita increase, cigarette consumption will increase as well, while real price of cigarette per pack increase, the cigarette consumption will decrease.

1.2) What is the elasticity of demand for cigarettes with respect to price? Is it statistically significant? If so, is it statistically different from 1?

$\ln C_i = 4.3 - 1.34 \ln P_i$

$\frac{\partial \ln C_i}{\partial \ln P_i} = \frac{\partial 4.3}{\partial \ln P_i} - \frac{\partial 1.34 \ln P_i}{\partial \ln P_i}$

$\frac{\partial \ln C_i}{\partial \ln P_i} = 1.34$

$\frac{C_i}{\partial C_i} \div \frac{P_i}{\partial P_i}$

$\frac{\partial C_i}{C_i} \left( \frac{P_i}{\partial P_i} \right)$

$\frac{\partial C_i}{\partial P_i} \times \frac{P_i}{C_i} = \varepsilon_p^C = 1.34$

$\varepsilon_p^c$ : Price Elasticity of Demand

- Test

$H_0 : \beta_2 = 0$

$H_a : \beta_2 \neq 0$

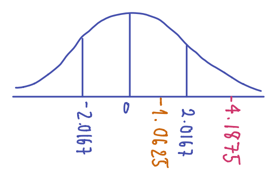
$t_{cal} = \frac{1.34 - 0}{0.32} = 4.1875$

$\alpha = 0.05, \; t_{\frac{0.05}{2}, 46-3} = 2.0167$

∴ We reject $H_0$. $\beta_2$ is statistically significant.

$H_0 : \beta_2 = 1$

$H_a : \beta_2 \neq 1$




$t_{cal} = \frac{1.34 - 1}{0.32} = 1.0625$

$t_{\frac{0.05}{2}, 46-3} = 2.0167$

∴ We fail to reject $H_0$. $\beta_2$ is statistically different from 1.

1.3) What is the income elasticity of demand for cigarettes? Is it statistically significant? If not, what might be the reasons for it?

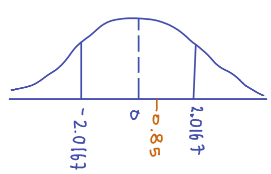
T-test

$H_0 : \beta_3 = 0$

$H_a : \beta_3 \neq 0$

$t_{cal} = \frac{0.17 - 0}{0.2} = 0.85$

$t_{\frac{0.05}{2}, 46-3} = 2.0167$




∴ We fail to reject $H_0$.

$\beta_3$ isn't statistically significant because sampling process might not be efficient. Also, real disposable income which real price and real income could be correlated.

6204641333

**2.** From estimating the regression equation on net financial wealth (nettfa), age of the survey respondent (age), and annual family income (inc) for people in the United States. The wealth and income variables are both recorded in thousands of dollars. The OLS estimation results for the model are given by

$$nettfa_i = \beta_1 + \beta_2 inc_i + \beta_3 age_i + u_i$$

```
reg nettfa inc age

      Source |       SS           df       MS      Number of obs   =     9,275
-------------+----------------------------------   F(2, 9272)      =    943.21
       Model |  6414618.8         2   3207309.4   Prob > F        =    0.0000
    Residual |  31528770.7     9,272  3400.42825   R-squared       =    0.1691
-------------+----------------------------------   Adj R-squared   =    0.1689
       Total |  37943389.5     9,274  4091.3726   Root MSE        =    58.313

------------------------------------------------------------------------------
      nettfa |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         inc |   .9533566   .0252775    37.72   0.000     .9038072    1.002906
         age |   1.030777   .0591226    17.43   0.000     .9148838     1.14667
       _cons |  -60.69654   2.596333   -23.38   0.000    -65.78592   -55.60715
------------------------------------------------------------------------------
```

```
reg nettfa inc age agesq

      Source |       SS           df       MS      Number of obs   =     9,275
-------------+----------------------------------   F(3, 9271)      =    646.80
       Model |  6567017.15         3  2189005.72   Prob > F        =    0.0000
    Residual |  31376372.3     9,271  3384.35685   R-squared       =    0.1731
-------------+----------------------------------   Adj R-squared   =    0.1728
       Total |  37943389.5     9,274  4091.3726   Root MSE        =    58.175

------------------------------------------------------------------------------
      nettfa |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         inc |   .9782522   .0254891    38.38   0.000      .928288    1.028216
         age |  -2.231489   .4897118    -4.56   0.000    -3.191432   -1.271547
       agesq |   .0377221   .0056214     6.71   0.000      .026703    .0487413
       _cons |   4.680388   10.08099     0.46   0.642    -15.08056    24.44134
------------------------------------------------------------------------------
```

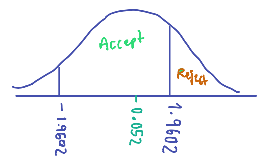
2.2) Due to estimation result by adding the age$^2$ variable or agesq. Perform the test whether we should include the quadratic term of the age variable or not? (Test for both t-test and F-test.) Also, interpret the meaning of this coefficient.

2.1) Test the coefficient, in the first model, $\beta_3 < 1$ in the first model or not?

T-test

$H_1 : \beta_3 < 1$

$H_1 : \beta_3 \nless 1$




$$t = \frac{1.030777 - 1}{0.0591226} = 0.5206$$

$$t_{\frac{\alpha}{2}, n-3} = t_{\frac{0.05}{2}, 9275-3} = t_{0.05, 9272} = 1.9602$$

We fail to reject null-hypothesis

$\therefore \beta_3 < 1$ is in the first model because it falls in region of acceptance.

2.2)

T-test

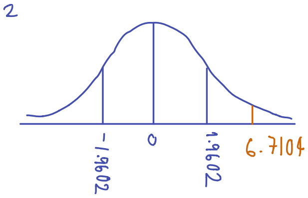
$H_0 : \beta_4 = 0$

$H_a$ : Otherwise

$$t_{cal} = \frac{0.0377221 - 0}{0.0056214} = 6.7104$$

$\alpha = 0.05$

$$t_{\frac{0.05}{2}, 9275-4} = 1.9602$$




$\therefore$ We reject $H_0$

$H_0$: $\beta_4$ has a contribution to the model

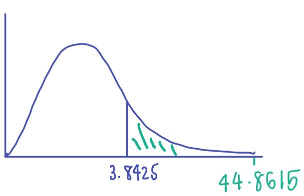
$H_a$ : otherwise

$$F_{cal} = \frac{R^2_{new} - R^2_{old} \,/\, \text{\# of new regression}}{1 - R^2_{new} \,/\, n - k_{new}}$$

$$= \frac{0.1731 - 0.1691 \,/\, 1}{1 - 0.1731 \,/\, 9275 - 1} = \frac{0.004}{0.00008916} = 44.8632$$

$\alpha = 0.05$

$F_{upper, \alpha}(1, 9271) = 3.8425$




$\therefore$ We reject $H_0$ because it falls in artical region. → $\beta_4$ doesn't contribute to model

Ans.

If we square age, it becomes a quadatric equation which instead of making the curve a straight line, it makes the curve U shape instead and squared age doesn't help explaining more about the model meaning that it doesn't help increasing ESS

6204641333

**3.** You are conducting an empirical investigation into the median prices of houses in 506 communities of a large metropolitan area. The sample data consist of 506 observations on the following observable variables:

- $P_i$: the median house price in community $i$, in dollars;
- $NOX_i$: the level of nitrous oxide in the air of community $i$, in parts per 100 million;
- $DIST_i$: the weighted distance of community $i$ from municipal area, in miles;
- $ROOM_i$: the average number of rooms per house in community $i$;
- $STRAT_i$: the average student-teacher ratio of schools in community $i$.

Researcher estimates the following model of median house price. The OLS estimation results for the model are given by

$$ln(P_i) = 11.08 - 0.9535\,ln(NOX_i) - 0.1343\,ln(DIST_i) + 0.2545ROOM_i - 0.05245STRAT_i$$
$$se \quad = (0.3181)\ (0.1167) \quad (0.04310) \quad (0.01853) \quad (0.005897)$$
$$RSS = 35.1835 \quad TSS = 84.5822$$

3.1) Interpret each of the coefficient estimates in regression equation.

If $NOX_i$ increases by 1 %, $P_i$ will decrease 0.9535 %.
If $DIST_i$ increases by 1 %, $P_i$ will decrease 0.1343 %.
If $ROOM_i$ increases by 1 %, $P_i$ will increase 0.2545 %.
If $STRAT_i$ increases by 1 %, $P_i$ will decrease 0.05245 %.

3.2) Test the individual significance of each of the slope coefficient estimates for $ln(NOX_i)$ and $ROOM_i$.

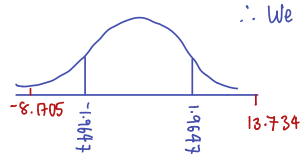
Test $\beta_2$

$H_0: \beta_2 = 0$
$H_a:$ otherwise
95 % CI

$t_{cal} = \dfrac{-0.9535 - 0}{0.1167} = -8.1705$

$t_{\frac{\alpha}{2}} = t_{\frac{0.05}{2},\,506-5} = 1.9647$

∴ We reject $H_0$, so, slope coefficient of ln(NOX) is statistically significant from 0.




$H_0: \beta_4 = 0$
$H_a:$ otherwise

$t = \dfrac{0.2545 - 0}{0.01853} = 13.734$

$t_{0.025,\,501} = 1.9647$

∴ We reject $H_0$ because slope coefficient of ROOM is statistically significant from 0.

3.3) Find the R-squared, adjusted R-squared, and test the joint significance of all the slope coefficient estimates.

$\bar{R}^2$

$R^2 = 1 - \dfrac{RSS}{TSS} = 1 - \dfrac{35.1835}{84.5822} = 0.584$

$\bar{R}^2 = 1 - (1-R^2)\dfrac{n-1}{n-k} = 1 - (1-0.584)\dfrac{506-1}{506-5} = 0.5807$

Test the joint significant

- $H_0: \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$

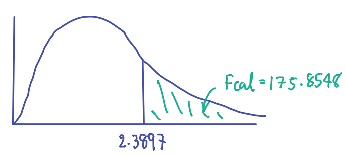
  $H_a:$ otherwise

$TSS = RSS + ESS$
$84.5822 = 35.1835 + ESS$
$ESS = 49.3987$

$F_{cal} = \dfrac{ESS}{RSS} \cdot \dfrac{(n-k)}{k-1} = \dfrac{49.3987}{35.1835} \cdot \dfrac{(506-5)}{(5-1)} = 175.8548$

$\alpha = 0.05$;

$F_{0.05}(4, 501) = 2.3897$




∴ We reject $H_0$. $F_{cal} > F_{table}$

→ Not all slope are simultaneously zero.

3.4) If researcher would like to test the proposition that the marginal effect of $ln(NOX_i)$ on $ln(P_i)$ equals the marginal effect of $ln(DIST_i)$ on $ln(P_i)$, write the restricted model and perform the test comparing restricted and unrestricted model, given that OLS estimation of this restricted regression equation yields a Residual Sum of Squares value = 41.9532.

restrict $\beta_2 = \beta_3$

$ln(P_i) = 11.08 - 0.9535\,ln(NOX_i) - 0.1343\,ln(DIST_i) + 0.2545\,ROOM_i - 0.05245\,STRAT_i$

$= 11.08 - 0.9535\,ln(NOX_i) - 0.9535\,ln(DIST_i) + 0.2545\,ROOM_i - 0.05245\,STRAT_i$

$- 0.9535[ln(NOX_i) + ln(DIST_i)]$

$= 11.08 - 0.9535\,(ln(NOX_i \cdot DIST_i)) + 0.2545\,ROOM_i - 0.05245\,STRAT_i$

Hypothesis

- $H_0: \beta_2 = \beta_3$

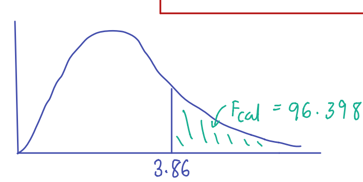
  $H_a:$ otherwise

$RSS_R = 41.9532$
$RSS_{UR} = 35.1835$

$\alpha = 0.05$
$F_{(1,501)} = 3.86$

$F_{cal} = \dfrac{RSS_R - RSS_{UR}/m}{RSS_{UR}/(n-k_{UR})}$

$= \dfrac{41.9532 - 35.1835 / 1}{35.1835 / 506 - 5} = 96.398$




∴ We reject $H_0$ because $F_{cal} > F_{table}$

The restriction imposed is not valid meaning that $\beta_2 \neq \beta_3$.

6204641333

**4.** Production function (Y) of the industrial sector in Thailand. It depends on the capital factor (K) and labor factor (L) in the years 1980-2010 with the following estimation.

**Model 1**:

$\ln Y_t = 18.27 + 0.536 \ln L_t + 0.024 \ln K_t$

$R^2 = 0.9389, RSS = 0.0124$

**Model 2**:

$\ln\left(\frac{Y}{L}\right)_t = 2.13 + 1.12 \ln\left(\frac{K}{L}\right)_t$

$R^2 = 0.8087, RSS = 0.0153$

4.1) Interpret the coefficients of the independent variables in models 1 and 2.

4.1)

Model 1 : If the capital factor increases by 1 %, the production function will increase 0.024 %
If the labor factor increases by 1 %, the production function will increase 0.536 %.

Model 2 : If capital per labor increases by 1 %, the output per labor will increase 1.12 %.

4.2) Test the hypothesis. Is the industrial production function characterized by constant return to scale? (Hint: you can perform any type of test that you see fit.)

Test the hypothesis

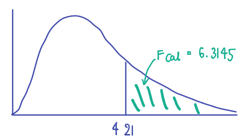
$H_0 : \beta_2 + \beta_3 = 1$

$H_a$ : otherwise

$F_{cal} = \frac{RSS_R - RSS_{UR} / m}{RSS_{UR} / (n - K_{UR})} = \frac{0.0153 - 0.0124 / 1}{0.0124 / 30 - 3} = 6.3145$

$\alpha = 0.05$

$F_{(1,27)} = 4.21$




∴ We reject $H_0$ because $F_{cal} > F_{table}$. The production function is not constant return to scale.

4.3) Can we compare the $R^2$ value between the two regression models? Why?

we cannot compare the $R^2$ value between two regression models because the left hand side of these two models (Model 1 : $\ln(Y_t)$ and Model 2 : $\ln\left(\frac{Y}{L}\right)_t$) are not the same.